



KESHAV MEMORIAL INSTITUTE OF COMMERCE & SCIENCES

Affiliated to Osmania University, NAAC Accredited

Narayanaguda, Hyd

A Unit of Keshav Memorial Educational Society

Report on

Guest Lecture

Data Pre Processing & Visualization

Session Details

Title of the session: Data Preprocessing & Visualization using Python

Date: 03/07/2026

Duration: 1 hour

Activity Category: IIC Calendar/MIC Driven/Self Driven/Celebration

Theme: To gain the knowledge on Data Preprocessing and visualization

Venue: Bsc Computer Science Lab,U.G Block

Organised by: Keshav Memorial Institute of Commerce and Sciences, Department of Computer Science

Expert/Speaker Details:

Name: Ms. C.Sanjana

Designation: Assistant Professor of English

Organization: Keshav Memorial Institute of Commerce & Sciences, Narayanaguda, Hyderabad.

Brief about Expert/Speaker: Ms. . C.Sanjana, Assistant professor at KMICS,Narayanaguda, Hyderabad. She is a dedicated academician currently pursuing a Ph.D in English. She Recognized nationally with the Shilpa Icon Award 2025, she is known for her contributions to literature, pedagogy, and student mentorship.

Outcome/Report of the activity:

Introduction

A guest lecture on "**Data Preprocessing and Visualization in Python**" was organized to enhance students' understanding of data analytics and machine learning workflows. The session aimed to introduce participants to the importance of preparing raw data before analysis and creating meaningful visualizations for better decision-making.

The lecture combined theoretical concepts with practical demonstrations using Python libraries, enabling students to understand real-world data processing techniques.

Objectives

The primary objectives of the guest lecture were:

- To understand the importance of data preprocessing in data analysis.
 - To learn techniques for cleaning and transforming datasets.
 - To introduce Python libraries used in data preprocessing.
 - To understand the role of data visualization in exploratory data analysis.
 - To demonstrate visualization techniques using Python.
 - To encourage students to apply preprocessing and visualization techniques in research and projects.
-

Topics Covered

1. Introduction to Data Science

- Data Science workflow
 - Importance of data quality
 - Applications of Python in Data Science
-

2. Data Preprocessing

The speaker explained that real-world datasets often contain errors, missing values, duplicates, and inconsistent formats. Proper preprocessing improves model accuracy and analysis quality.

The following preprocessing techniques were demonstrated:

- Importing datasets using Pandas
 - Understanding dataset structure
 - Handling missing values
 - Removing duplicate records
 - Data cleaning
 - Data transformation
 - Feature scaling
 - Encoding categorical variables
 - Detecting and treating outliers
 - Data normalization and standardization
-

3. Python Libraries Used

The resource person introduced several popular Python libraries including:

- Pandas
- NumPy
- Matplotlib
- Plotly
- Seaborn
- Scikit-learn

Their applications in preprocessing and visualization were explained with examples.

4. Exploratory Data Analysis (EDA)

Students learned how to perform exploratory analysis by:

- Viewing summary statistics
 - Checking data types
 - Understanding variable distributions
 - Correlation analysis
 - Identifying trends and patterns
-

5. Data Visualization Techniques

Different visualization methods were demonstrated to present data effectively.

These included:

- Line Charts
- Bar Charts
- Pie Charts
- Histograms
- Scatter Plots
- Box Plots
- Heat Maps
- Pair Plots
- Distribution Plots

The speaker emphasized selecting the appropriate chart based on the nature of the data.

6. Hands-on Demonstration

A practical session was conducted where students learned to:

- Load datasets into Python
- Clean and preprocess data

- Handle missing values
 - Create statistical summaries
 - Generate attractive visualizations
 - Interpret graphical outputs
-

Learning Outcomes

After attending the lecture, participants were able to:

- Understand the complete data preprocessing workflow.
 - Clean and prepare datasets for analysis.
 - Handle missing and inconsistent data.
 - Perform exploratory data analysis using Python.
 - Create informative visualizations.
 - Interpret graphs for data-driven decision-making.
 - Apply preprocessing techniques in machine learning projects.
-

Benefits to Students

The lecture helped students to:

- Gain practical exposure to Python programming.
 - Improve analytical thinking.
 - Understand industry-standard data preprocessing techniques.
 - Learn visualization methods used in research and business analytics.
 - Build confidence in handling real-world datasets.
-

Interactive Session

An interactive question-and-answer session was conducted at the end of the lecture. Students actively participated by asking questions related to:

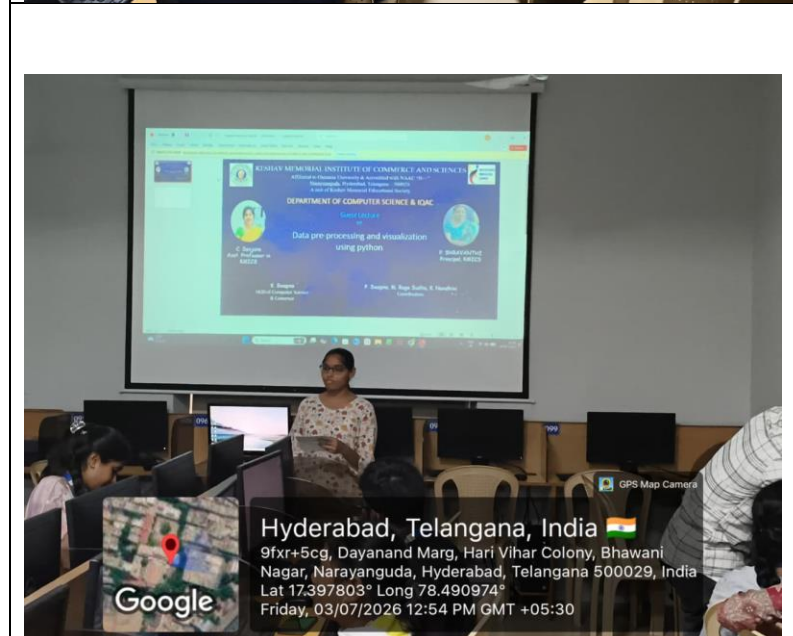
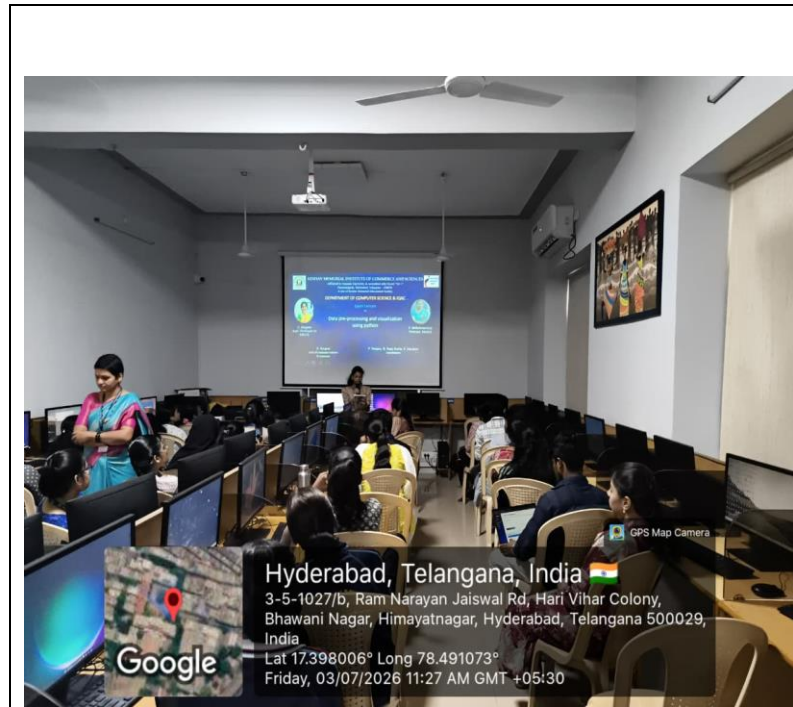
- Handling large datasets
- Missing value imputation techniques
- Choosing appropriate visualization methods
- Python libraries for data analysis
- Career opportunities in Data Science and Machine Learning

The resource person addressed all the queries with practical examples.

Participant details:

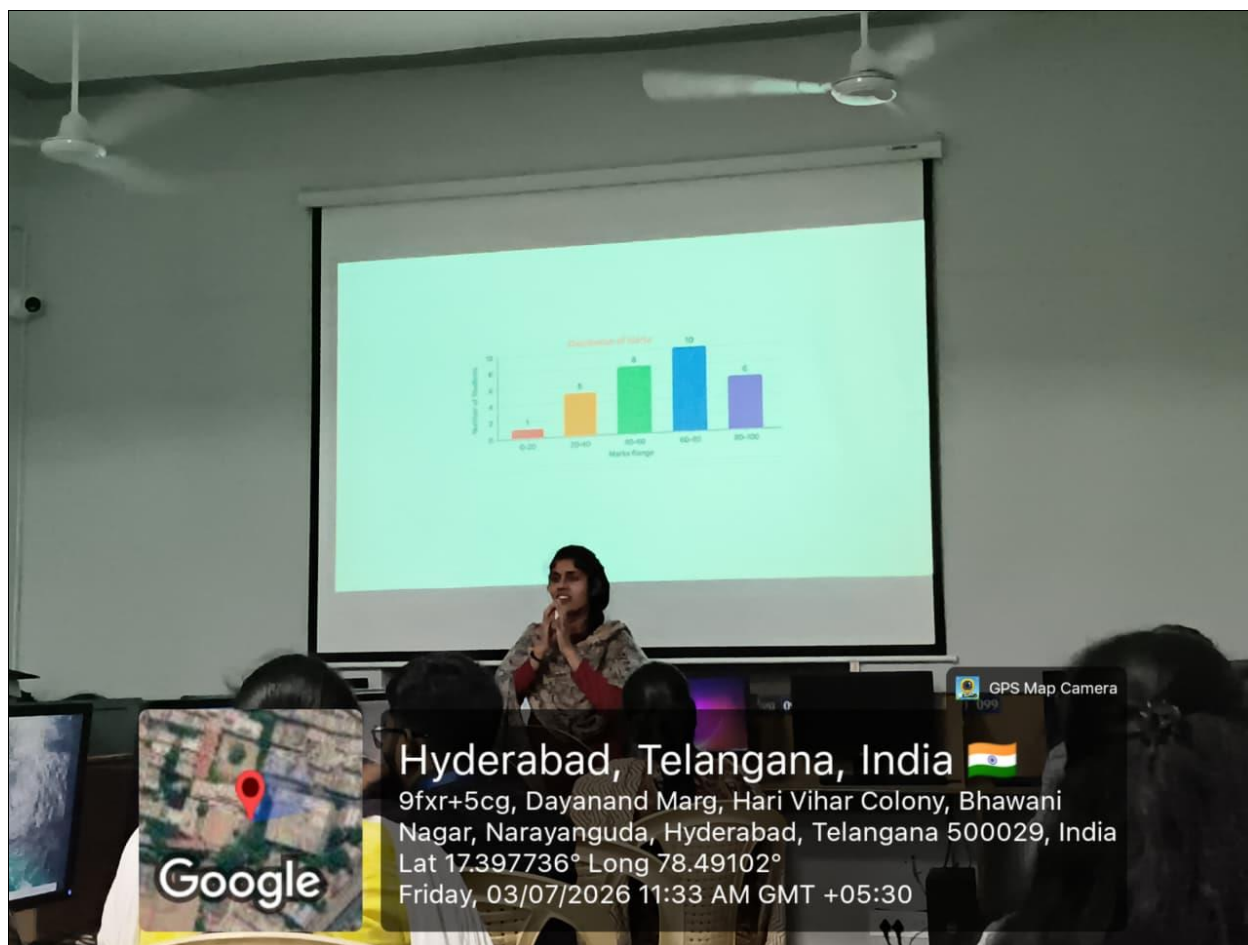
Total no. of Student participation: **60**(MSc Computer Science I & II Year)

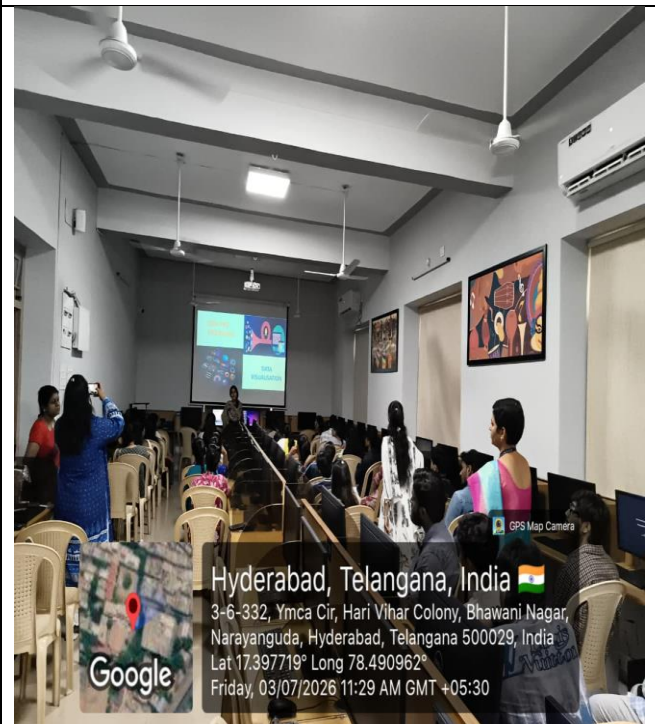
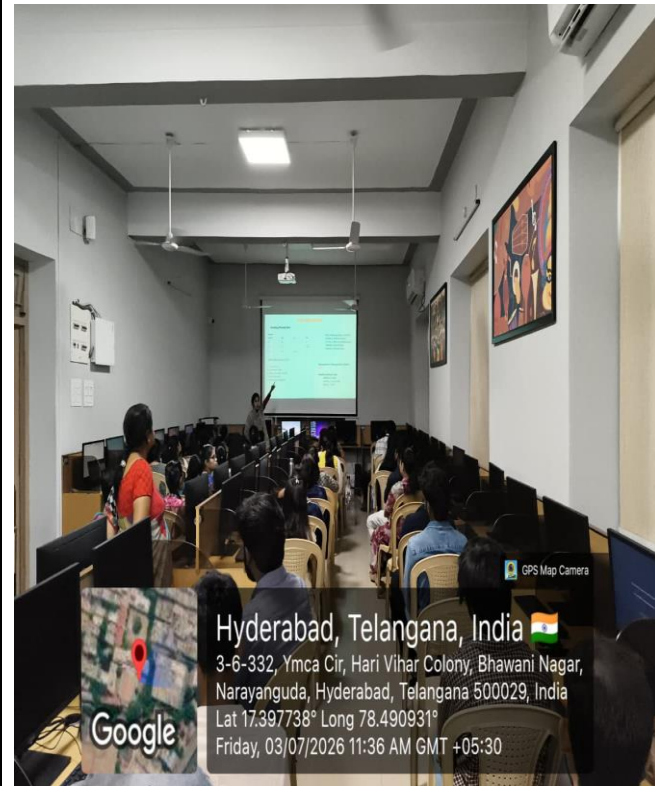
Total no. of Staff (Teaching/Non-teaching) participation: 4



- Attendance details /Participants List (photo)(Word format provided)

Class	Year	No. of Students
MSc Computer Science	I Year	20
MSc Computer Science	II Year	40





Conclusion

The guest lecture on "**Data Preprocessing and Visualization in Python**" was highly informative and beneficial. It provided participants with both theoretical knowledge and practical skills required for effective data analysis. The hands-on demonstrations helped students understand the significance of preprocessing and visualization in extracting meaningful insights from data.

Overall, the session successfully enhanced students' understanding of Python-based data analytics and motivated them to apply these techniques in academic research, machine learning projects, and future professional careers.